

48  
DNT  
2-23-9

1647

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/005,318C

DATE: 02/06/2001  
TIME: 15:57:09

Input Set: A:\401c1.app  
Output Set: N:\CRF3\02062001\I005318C.raw

ENTERED

RECEIVED

FEB 20 2001

TECH CENTER 1600/2900

## SEQUENCE LISTING

4 (1) GENERAL INFORMATION:  
C--> 6 (i) APPLICANT: Hein, Mich B.  
7 Hiatt, Andrew C.  
8 Fitchen, John H.  
10 (ii) TITLE OF INVENTION: NOVEL EPITHELIAL TISSUE TARGETING AGENT  
12 (iii) NUMBER OF SEQUENCES: 140  
14 (iv) CORRESPONDENCE ADDRESS:  
15 (A) ADDRESSEE: SEED IP LAW GROUP PLLC  
16 (B) STREET: 701 Fifth Avenue, Suite 6300  
17 (C) CITY: Seattle  
18 (D) STATE: Washington  
19 (E) COUNTRY: USA  
20 (F) ZIP: 98104  
22 (v) COMPUTER READABLE FORM:  
23 (A) MEDIUM TYPE: Floppy disk  
24 (B) COMPUTER: IBM PC compatible  
25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30  
28 (vi) CURRENT APPLICATION DATA:  
C--> 29 (A) APPLICATION NUMBER: US/09/005,318C  
C--> 30 (B) FILING DATE: 09-Jan-1998  
31 (C) CLASSIFICATION:  
33 (viii) ATTORNEY/AGENT INFORMATION:  
34 (A) NAME: Christiansen, William T.  
35 (B) REGISTRATION NUMBER: 44,614  
36 (C) REFERENCE/DOCKET NUMBER: 310098.401C1  
38 (ix) TELECOMMUNICATION INFORMATION:  
39 (A) TELEPHONE: (206) 622-4900  
40 (B) TELEFAX: (206) 682-6031  
43 (2) INFORMATION FOR SEQ ID NO: 1:  
45 (i) SEQUENCE CHARACTERISTICS:  
46 (A) LENGTH: 137 amino acids  
47 (B) TYPE: amino acid  
48 (C) STRANDEDNESS:  
49 (D) TOPOLOGY: linear  
55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
57 Gln Glu Asp Glu Arg Ile Val Leu Val Asp Asn Lys Cys Lys Cys Ala  
58 1 5 10 15  
60 Arg Ile Thr Ser Arg Ile Ile Arg Ser Ser Glu Asp Pro Asn Glu Asp  
61 20 25 30  
63 Ile Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Asn Arg Glu  
64 35 40 45  
66 Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Arg Pro Val Tyr His  
67 50 55 60  
69 Leu Ser Asp Leu Cys Lys Lys Cys Asp Pro Thr Glu Val Glu Leu Asp  
70 65 70 75 80

RAW SEQUENCE LISTING      DATE: 02/06/2001  
 PATENT APPLICATION: US/09/005,318C      TIME: 15:57:09

Input Set : A:\401c1.app  
 Output Set: N:\CRF3\02062001\I005318C.raw

```

72 Asn Gln Ile Val Thr Ala Thr Gln Ser Asn Ile Cys Asp Glu Asp Ser
73                               85          90          95
75 Ala Thr Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Ala
76                               100         105         110
78 Val Val Pro Leu Val Tyr Gly Gly Glu Thr Lys Met Val Glu Thr Ala
79                               115         120         125
81 Leu Thr Pro Asp Ala Cys Tyr Pro Asp
82   130          135
84 (2) INFORMATION FOR SEQ ID NO: 2:
86   (i) SEQUENCE CHARACTERISTICS:
87       (A) LENGTH: 135 amino acids
88       (B) TYPE: amino acid
89       (C) STRANDEDNESS:
90       (D) TOPOLOGY: linear
96   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
98   Gln Asp Glu Asn Glu Arg Ile Val Val Asp Asn Lys Cys Lys Cys Ala
99   1           5           10          15
101  Arg Ile Thr Ser Arg Ile Ile Pro Ser Ala Glu Asp Pro Ser Gln Asp
102           20          25          30
104  Ile Val Glu Arg Asn Val Arg Ile Ile Val Pro Leu Asn Ser Arg Glu
105           35          40          45
107  Asn Ile Ser Asp Pro Thr Ser Pro Met Arg Thr Lys Pro Val Tyr His
108           50          55          60
110  Leu Ser Asp Leu Cys Lys Lys Cys Asp Thr Thr Glu Val Glu Leu Glu
111           65          70          75          80
113  Asp Gln Val Val Thr Ala Ser Gln Ser Asn Ile Cys Asp Ser Asp Ala
114           85          90          95
116  Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Asn Arg Val
117           100         105         110
119  Lys Leu Ser Tyr Arg Gly Gln Thr Lys Met Val Glu Thr Ala Leu Thr
120           115         120         125
122  Pro Asp Ser Cys Tyr Pro Asp
123   130          135
126 (2) INFORMATION FOR SEQ ID NO: 3:
128   (i) SEQUENCE CHARACTERISTICS:
129       (A) LENGTH: 137 amino acids
130       (B) TYPE: amino acid
131       (C) STRANDEDNESS:
132       (D) TOPOLOGY: linear
138   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
140  Asp Asp Glu Ala Thr Ile Leu Ala Asp Asn Lys Cys Met Cys Thr Arg
141   1           5           10          15
143  Val Thr Ser Arg Ile Ile Pro Ser Thr Glu Asp Pro Asn Glu Asp Ile
144           20          25          30
146  Val Glu Arg Asn Ile Arg Ile Val Val Pro Leu Asn Asn Arg Glu Asn
147           35          40          45
149  Ile Ser Asp Pro Thr Ser Pro Leu Arg Arg Asn Pro Val Tyr His Leu
150           50          55          60
152  Ser Asp Val Cys Lys Lys Cys Asp Pro Val Glu Val Glu Leu Glu Asp

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/005,318C

DATE: 02/06/2001

TIME: 15:57:09

Input Set : A:\401c1.app

Output Set: N:\CRF3\02062001\I005318C.raw

**RECEIVED****FEB 20 2001****DATA CENTER 1600/2800**

```

153      65      70      75      80
155      Gln Val Val Thr Ala Thr Gln Ser Asn Ile Cys Asn Glu Asp Asp Gly
156      85      90      95
158      Val Pro Glu Thr Cys Tyr Met Tyr Asp Arg Asn Lys Cys Tyr Thr Thr
159      100     105     110
161      Met Val Pro Leu Arg Tyr His Gly Glu Thr Lys Met Val Gln Ala Ala
162      115     120     125
164      Leu Thr Pro Asp Ser Cys Tyr Pro Asp
165      130     135
167 (2) INFORMATION FOR SEQ ID NO: 4:
169      (i) SEQUENCE CHARACTERISTICS:
170          (A) LENGTH: 136 amino acids
171          (B) TYPE: amino acid
172          (C) STRANDEDNESS:
173          (D) TOPOLOGY: linear
179      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
181      Glu Asp Glu Ser Thr Val Leu Val Asp Asn Lys Cys Gln Cys Val Arg
182      1      5      10      15
184      Ile Thr Ser Arg Ile Ile Arg Asp Pro Asp Asn Pro Ser Glu Asp Ile
185      20     25     30
187      Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Thr Arg Glu Asn
188      35     40     45
190      Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Glu Pro Lys Tyr Asn Leu
191      50     55     60
193      Ala Asn Leu Cys Lys Lys Cys Asp Pro Thr Glu Ile Glu Leu Asp Asn
194      65     70     75     80
196      Gln Val Phe Thr Ala Ser Gln Ser Asn Ile Cys Pro Asp Asp Asp Tyr
197      85     90     95
199      Ser Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Thr Leu
200      100    105    110
202      Val Pro Ile Thr His Arg Gly Val Thr Arg Met Val Lys Ala Thr Leu
203      115    120    125
205      Thr Pro Asp Ser Cys Tyr Pro Asp
206      130    135
208 (2) INFORMATION FOR SEQ ID NO: 5:
210      (i) SEQUENCE CHARACTERISTICS:
211          (A) LENGTH: 119 amino acids
212          (B) TYPE: amino acid
213          (C) STRANDEDNESS:
214          (D) TOPOLOGY: linear
220      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
222      Glu Gln Glu Tyr Ile Leu Ala Asn Asn Lys Cys Lys Cys Val Lys Ile
223      1      5      10      15
225      Ser Ser Arg Phe Val Pro Ser Thr Glu Arg Pro Gly Glu Glu Ile Leu
226      20     25     30
W--> 228      Glu Arg Asn Ile Gln Ile Thr Ile Pro Thr Ser Ser Arg Met Xaa Ile
229      35     40     45
231      Ser Asp Pro Tyr Ser Pro Leu Arg Thr Gln Pro Val Tyr Asn Leu Trp
232      50     55     60

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/005,318C

DATE: 02/06/2001

TIME: 15:57:09

Input Set : A:\401c1.app

Output Set: N:\CRF3\02062001\I005318C.raw

```

234   Asp Ile Cys Gln Lys Cys Asp Pro Val Gln Leu Glu Ile Gly Gly Ile
235   65              70              75              80
W--> 237   Pro Val Leu Ala Ser Gln Pro Xaa Xaa Ser Xaa Pro Asp Asp Glu Cys
238   85              90              95
240   Tyr Thr Thr Glu Val Asn Phe Lys Lys Lys Val Pro Leu Thr Pro Asp
241   100             105             110
243   Ser Cys Tyr Glu Tyr Ser Glu
244   115
247 (2) INFORMATION FOR SEQ ID NO: 6:
249   (i) SEQUENCE CHARACTERISTICS:
250       (A) LENGTH: 128 amino acids
251       (B) TYPE: amino acid
252       (C) STRANDEDNESS:
253       (D) TOPOLOGY: linear
259   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
261   Asn Lys Cys Met Cys Thr Arg Val Thr Ala Arg Ile Arg Gly Thr Arg
262   1              5              10              15
264   Glu Asp Pro Asn Glu Asp Ile Val Glu Arg Tyr Ile Arg Ile Asn Val
265   20             25             30
267   Pro Leu Lys Asn Arg Gly Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg
268   35             40             45
270   Asn Gln Pro Val Tyr His Leu Ser Pro Ser Cys Lys Lys Cys Asp Pro
271   50             55             60
273   Tyr Glu Asp Gly Val Val Thr Ala Thr Glu Thr Asn Ile Cys Tyr Pro
274   65             70             75             80
276   Asp Glu Gly Val Pro Gln Ser Cys Arg Asp Tyr Cys Pro Glu Leu Asp
277   85             90             95
279   Arg Asn Lys Cys Tyr Thr Val Leu Val Pro Pro Gly Tyr Thr Gly Glu
280   100            105            110
282   Thr Lys Met Val Gln Asn Ala Leu Thr Pro Asp Ala Cys Tyr Pro Asp
283   115            120            125
286 (2) INFORMATION FOR SEQ ID NO: 7:
288   (i) SEQUENCE CHARACTERISTICS:
289       (A) LENGTH: 421 base pairs
290       (B) TYPE: nucleic acid
291       (C) STRANDEDNESS: single
292       (D) TOPOLOGY: linear
296   (ix) FEATURE:
297       (A) NAME/KEY: CDS
298       (B) LOCATION: 1..414
301   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
303   GAT CAG GAA GAT GAA CGT ATT GTT CTG GTT GAC AAC AAG TGC AAG TGT      48
304   Asp Gln Glu Asp Glu Arg Ile Val Leu Val Asp Asn Lys Cys Lys Cys
305   1              5              10              15
307   GCT CGT ATT ACT TCT AGA ATC ATC CGT AGC TCA GAG GAC CCA AAT GAA      96
308   Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser Ser Glu Asp Pro Asn Glu
309   20             25             30
311   GAT ATA GTC GAA CGT AAC ATC CGT ATC ATC GTC CCA CTG AAT AAC CGG      144
312   Asp Ile Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Asn Arg

```

RAW SEQUENCE LISTING      DATE: 02/06/2001  
 PATENT APPLICATION: US/09/005,318C      TIME: 15:57:09

Input Set : A:\401c1.app  
 Output Set : N:\CRF3\02062001\I005318C.raw

```

313          35          40          45
315 GAG AAT ATC TCA GAT CCT ACA AGT CCG TTG CGC ACA CGC TTC GTA TAC      192
316 Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Arg Phe Val Tyr
317          50          55          60
319 CAC CTG TCA GAT CTG TGT AAG AAG TGT GAT CCA ACA GAG GTA GAG CTG      240
320 His Leu Ser Asp Leu Cys Lys Lys Cys Asp Pro Thr Glu Val Glu Leu
321 65          70          75          80
323 GAC AAT CAG ATA GTC ACT GCG ACT CAA AGC AAC ATT TGC GAT GAG GAC      288
324 Asp Asn Gln Ile Val Thr Ala Thr Gln Ser Asn Ile Cys Asp Glu Asp
325          85          90          95
327 AGC GCT ACA GAA ACC TGC AGC ACC TAC GAT AGG AAC AAA TGC TAC ACG      336
328 Ser Ala Thr Glu Thr Cys Ser Thr Tyr Asp Arg Asn Lys Cys Tyr Thr
329          100          105          110
331 GCC GTG GTT CCG CTC GTG TAT GGT GGA GAG ACA AAA ATG GTG GAA ACT      384
332 Ala Val Val Pro Leu Val Tyr Gly Gly Glu Thr Lys Met Val Glu Thr
333          115          120          125
335 GCC CTT ACG CCC GAT GCA TGC TAT CCG GAC TGAATTC      421
336 Ala Leu Thr Pro Asp Ala Cys Tyr Pro Asp
337          130          135
340 (2) INFORMATION FOR SEQ ID NO: 8:
342 (i) SEQUENCE CHARACTERISTICS:
343 (A) LENGTH: 215 base pairs
344 (B) TYPE: nucleic acid
345 (C) STRANDEDNESS: single
346 (D) TOPOLOGY: linear
350 (ix) FEATURE:
351 (A) NAME/KEY: CDS
352 (B) LOCATION: 1..213
355 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
357 GAT CAG AAG TGC AAG TGT GCT CGT ATT ACT TCT AGA ATC ATC CGT AGC      48
358 Asp Gln Lys Cys Lys Cys Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser
359 1          5          10          15
361 TCA GAG GAC CCA AAT GAA GAT ATA GTC GAA CGT AAC ATC CGT ATC ATC      96
362 Ser Glu Asp Pro Asn Glu Asp Ile Val Glu Arg Asn Ile Arg Ile Ile
363          20          25          30
365 GTC CCA CTG AAT AAC CGG GAG AAT ATC TCA GAT CCT ACA AGT CCG TTG      144
366 Val Pro Leu Asn Asn Arg Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu
367          35          40          45
369 CGC ACA CGC TTC GTA TAC CAC CTG TCA GAT CTG TGT AAG AAG GAT GAG      192
370 Arg Thr Arg Phe Val Tyr His Leu Ser Asp Leu Cys Lys Lys Asp Glu
371          50          55          60
373 GAC AGC GCT ACA GAA ACC TGC TG      215
374 Asp Ser Ala Thr Glu Thr Cys
375 65          70
378 (2) INFORMATION FOR SEQ ID NO: 9:
380 (i) SEQUENCE CHARACTERISTICS:
381 (A) LENGTH: 140 base pairs
382 (B) TYPE: nucleic acid
383 (C) STRANDEDNESS: single

```

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/005,318C

DATE: 02/06/2001

TIME: 15:57:10

Input Set : A:\401c1.app

Output Set: N:\CRF3\02062001\I005318C.raw

L:6 M:220 C: Keyword misspelled or invalid format, [(i) APPLICANT:]  
L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:228 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:237 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:2470 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:118  
L:2473 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:118  
L:2624 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:123  
L:2630 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:123  
L:2633 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:123